

Vos données sont-elles réellement prêtes pour les agents IA ?

Des résultats IA fiables commencent par la qualité des données, pas par les modèles. Découvrez comment l'observabilité des données rend l'IA évolutive.



Table des matières

- 2 Résumé exécutif
- 3 Lorsque l'IA fonctionne avec des données défectueuses, chaque résultat est à risque
- 3 Ce que signifie réellement « données prêtes pour l'IA » en pratique
- 4 5 étapes pour construire une feuille de route pragmatique, centrée sur l'observabilité, à des données prêtes à l'IA
 - 4 1. Commencer par un cas d'usage à forte valeur ajoutée et à fort risque
 - 4 2. Faire de la fiabilité des données un processus continu, et non réactif
 - 4 3. Automatiser le triage et la résolution des incidents pour réduire le travail manuel
 - 4 4. Couvrir les formats de données nécessaires à vos systèmes d'IA
 - 4 5. Intégrer les contrôles de santé des données dans le workflow IA avec MCP
- 5 Comment Actian Data Observability propulse l'IA à l'échelle
- 5 Faire de la fiabilité des données un principe non négociable
- 5 À propos d'Actian

Résumé exécutif

La plupart des organisations ont déjà lancé des initiatives d'IA. Chatbots, copilotes, moteurs de recommandation et premiers agents IA font désormais partie des workflows quotidiens. Si l'enthousiasme est réel, de nombreuses équipes peinent encore à faire passer l'IA du stade pilote à des cas d'usage en production réellement fiables et dignes de confiance.

Selon Gartner, d'ici 2026, les organisations abandonneront 60 % des projets d'IA qui ne reposent pas sur des données bien préparées. Ce n'est pas un problème de choix de modèle ou de qualité d'algorithme, mais de fiabilité des données qui alimentent ces systèmes.

Contrairement aux analystes humains, les systèmes d'IA ne remettent pas en question les valeurs manquantes, les changements silencieux de schéma ou les retards dans les données. Ils supposent que les entrées sont valides et agissent en conséquence. Si les données sont incomplètes, obsolètes, biaisées ou mal surveillées, les résultats refléteront ces défauts, souvent à grande échelle et avec davantage d'assurance que les systèmes traditionnels.

Ce guide repose sur une vérité simple : **la préparation à l'IA commence par l'observabilité des données.**

Vous découvrirez ce qui se passe lorsque l'IA fonctionne avec des données défectueuses, à quoi ressemblent concrètement des données prêtes pour l'IA, pourquoi les agents IA renforcent la nécessité d'une fiabilité des données en temps réel et comment l'observabilité des données aide les organisations à détecter, expliquer et prévenir les incidents avant qu'ils n'impactent les analyses ou les systèmes d'IA. Vous verrez également comment des résultats insuffisants sont souvent directement liés à des lacunes en matière d'observabilité des données.

Alors, à quoi ressemble l'échec dans le monde réel ?

Lorsque l'IA fonctionne avec des données défectueuses, chaque résultat est à risque

Si l'IA repose sur des données erronées, les modes d'échec sont prévisibles et coûteux :

Des réponses incorrectes à grande échelle Un assistant IA orienté client s'appuie sur des tarifs obsolètes dans des milliers d'interactions avant que le problème ne soit détecté. Les équipes humaines doivent intervenir pour corriger les erreurs, ce qui érode la confiance envers l'IA et l'organisation.

Biais et distorsions invisibles Des données d'entraînement incomplètes ou déséquilibrées entraînent des décisions biaisées en matière de crédit ou de gestion des risques, souvent sans signaux d'alerte évidents jusqu'à ce que les régulateurs ou les clients soulèvent le problème.

Défaillances opérationnelles Des modèles de prévision, d'optimisation et de planification entraînés sur des données incohérentes allouent mal les stocks, évaluent incorrectement la demande ou fixent des prix inadaptés, avec un impact direct sur les revenus et l'expérience client.

Risques de conformité et d'audit Lorsque les résultats de l'IA ne peuvent être reliés à des sources de données validées et documentées, les audits deviennent difficiles et l'exposition réglementaire augmente.

Dans de nombreux cas, les modèles fonctionnent exactement comme prévu. Ce qui échoue, c'est l'hypothèse selon laquelle les données en amont restent correctes, complètes et à jour.

Ce que signifie réellement « données prêtes pour l'IA » en pratique

La préparation à l'IA est souvent présentée comme un problème d'outillage ou d'infrastructure. En réalité, il s'agit d'un problème de fiabilité des données.

Des données prêtes pour l'IA sont des données pouvant être utilisées en toute sécurité pour la prise de décision ou l'automatisation, car leur fraîcheur, leur complétude et leur comportement sont validés en continu.

Concrètement, cela signifie des données présentant les caractéristiques suivantes :

Fiables et à jour Les données sont exactes, complètes et actualisées. Les valeurs manquantes, les doublons, les changements de schéma et les variations de volume sont détectés avant d'impacter les modèles.

Observables en continu Les équipes peuvent visualiser le comportement des données dans le temps, y compris les anomalies, les dérives et les ruptures de pipeline, sans dépendre de contrôles manuels ponctuels.

Explicables en cas de changement Lorsque des incidents surviennent, les équipes comprennent rapidement ce qui a changé, où le changement s'est produit et pourquoi cela a un impact.

Accessibles sans approximation Les humains comme les systèmes d'IA peuvent vérifier si les données sont adaptées à l'usage avant de s'y appuyer.

Sans ces caractéristiques, les équipes restent souvent bloquées au stade du prototype, capables de créer des démonstrations d'IA, mais incapables de les déployer à grande échelle en production de manière fiable. Cette limite devient dangereuse lorsque les systèmes passent de l'aide à la décision à la prise de décision autonome. L'erreur la plus fréquente consiste à considérer la préparation à l'IA comme une certification ponctuelle plutôt qu'une exigence de validation continue.

5 étapes pour construire une feuille de route pragmatique, centrée sur l'observabilité, à des données prêtes à l'IA

Les organisations n'ont pas besoin de tout transformer d'un coup. La préparation à l'IA progresse plus rapidement lorsque l'observabilité des données est mise en œuvre de manière ciblée et progressive.

1. Commencer par un cas d'usage à forte valeur ajoutée et à fort risque

Les cas d'usage IA à haut risque incluent la souscription, la détermination d'éligibilité, le reporting réglementaire et les actions automatisées auprès des clients. Pour ces cas d'usage :

- Identifier les jeux de données critiques
- Définir des exigences minimales en matière de fraîcheur et de complétude
- Documenter les actions qui dépendent des résultats de l'IA

Ces étapes établissent une cible claire pour vos efforts en matière d'observabilité.

2. Faire de la fiabilité des données un processus continu, et non réactif

Les approches traditionnelles de qualité des données reposent sur des tableaux de bord et des contrôles hebdomadaires, qui révèlent souvent les problèmes uniquement après la dégradation des résultats de l'IA ou l'impact sur les clients. L'IA exige une validation continue dès l'ingestion des données. L'observabilité des données permet aux équipes de :

- Surveiller les pipelines et les jeux de données en temps réel
- Détecter les anomalies de manière proactive
- Identifier les dérives de données avant qu'elles ne dégradent les performances de l'IA

3. Automatiser le triage et la résolution des incidents pour réduire le travail manuel

Avant l'usage des agents, les équipes poursuivaient les alertes à travers différents outils et journaux. Avec les agents d'observabilité des données, les incidents sont validés, expliqués et orientés automatiquement. Ces agents prennent en charge l'ensemble du cycle de vie des problèmes de qualité des données :

- L'agent de validation contrôle les données dès leur ingestion
- L'agent de diagnostic des incidents explique les causes racines en langage clair
- L'agent de lignage montre l'impact en amont et en aval
- L'agent Insight met en évidence les tendances et les risques
- Les agents d'orchestration et de routage coordonnent la résolution
- L'agent d'assistance répond aux questions sur l'état de santé des données en langage naturel

Au lieu de courir après les alertes, les équipes bénéficient d'explications claires et d'une résolution guidée.

4. Couvrir les formats de données nécessaires à vos systèmes d'IA

La plupart des outils d'observabilité se concentrent sur les données structurées, telles que les tables de bases de données relationnelles, les tables de data warehouse ou de lakehouse, les feuilles de calcul et les séries temporelles. Pourtant, les pipelines IA réels reposent aussi sur des PDF, des fichiers XML, des contrats, des factures, des réclamations, des journaux et des flux partenaires, rarement supervisés avec la même rigueur que les tables.

Il est essentiel d'étendre la fiabilité de ces données aux formats. Cela offre aux équipes une visibilité sur les pipelines semi-structurés et documentaires, généralement hors du périmètre des outils de supervision traditionnels.

5. Intégrer les contrôles de santé des données dans le workflow IA avec MCP

Les systèmes d'IA ont besoin d'un moyen sûr de vérifier l'état de santé des données avant d'agir. Un serveur Model Context Protocol (MCP) expose facilement l'état de la qualité des données, les incidents, les alertes, les contrôles et les résultats de validation directement aux assistants IA et aux workflows agentiques.

Il permet aux systèmes d'IA de vérifier la santé des données avant de prendre une décision, plutôt que de se fier à des hypothèses ou à des signaux obsolètes.

Voici maintenant à quoi ressemble l'opérationnalisation de l'observabilité avec Actian.

Comment Actian Data Observability propulse l'IA à l'échelle

Actian Data Observability vise à rendre la fiabilité des données automatique et exploitable, avec une approche conçue pour les équipes qui déploient l'IA dans des environnements réglementés et à forts enjeux, où la confiance, l'explicabilité et la rapidité sont essentielles (Figure 1).

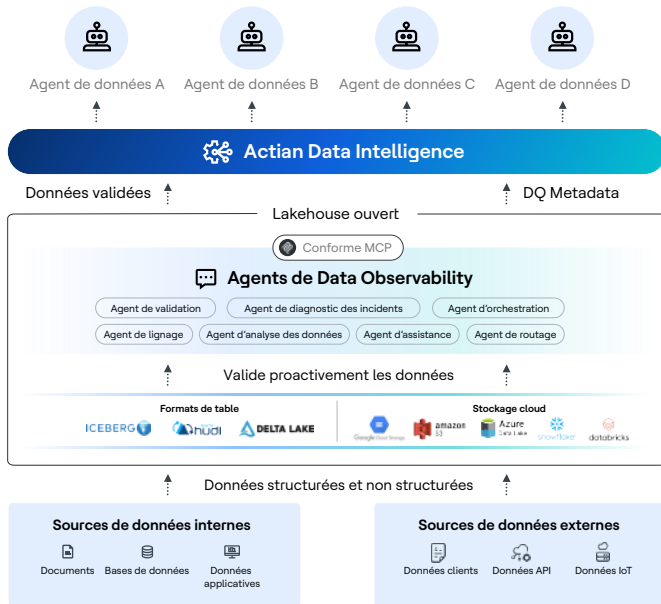


Figure 1 : Utiliser les agents pour l'observabilité des données

La solution Actian offre :

- Une validation continue des données provenant d'Iceberg, Delta Lake et Hudi.
- Des agents d'observabilité des données qui détectent, expliquent et résolvent rapidement les incidents.
- La prise en charge des données structurées, semi-structurées et documentaires.
- Une connectivité MCP qui apporte directement le contexte d'observabilité dans les outils IA.

Ensemble, ces capacités réduisent les risques, accélèrent la résolution et permettent aux équipes de faire évoluer l'IA sans ajouter de complexité ni d'effectifs supplémentaires.

Faire de la fiabilité des données un principe non négociable

La plupart des échecs liés à l'IA ne sont pas des échecs de modèles. Ce sont des défaillances de fiabilité des données qui apparaissent trop tard.

Si les résultats de l'IA sont incohérents, difficiles à expliquer ou peu fiables, la bonne question n'est pas « Quel modèle devons-nous utiliser ? », mais plutôt « Pouvons-nous vérifier que les données sont fiables à cet instant précis ? »

Actian Data Observability vous aide à répondre à cette question en continu, avant que les données n'atteignent les dirigeants, les clients ou des systèmes autonomes.

Si votre système d'IA ne peut pas vérifier l'état de santé des données au moment de l'action, il n'est pas prêt pour des décisions autonomes. Avec Actian Data Observability, vos équipes avancent plus vite en toute confiance, en faisant de la santé des données un contrôle intégré qui garantit des décisions IA fiables, explicables et prêtes à passer à l'échelle.

À propos d'Actian

Actian aide les entreprises à gérer et gouverner leurs données en toute confiance, à grande échelle. Les organisations font confiance aux solutions de gestion et d'intelligence des données d'Actian pour maîtriser des environnements data complexes et accélérer la mise à disposition de données prêtes pour l'IA. Conçues pour être flexibles et fiables, ces solutions s'intègrent facilement à tous types d'environnements : sur site, dans le cloud ou en mode hybride. Actian est la division Data & IA de HCLSoftware. Pour en savoir plus : www.actian.com.